

WE CLAIM:

1. A nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of SEQ ID NOS: 4-6, 8, 12-14, 16-27, 30-50, 52, 53, 55-69, 72-86, 88-92, 97-100, 102-121, 123, 125, 127-144, 146-152, 154-164, 167, 168, 171-173, 175-182, 184-191, 193, 194, 196-205, 207, 209-214, 216-219, 221-225, 227-245, 250-252, 254, 256-260, 264-270, 273-281, 285-292, 294, 296-300, 303, 304, 306-313, 315-318, 320, 322-329, 331-335, 339-352, 354, 355, 357-388, 391, 393-426, 428, 429, 432-468, 470-486, 488-495, 497-507, 509-513, 515-525, 527, 528, 532, 533, 535, 538, 540-542, 545-556, 558-563, 565-576, 578-597, 599-618, 620, 624-628, 630, 631, 637-656, 659-662, 665, 667-675, 677-683, 685-697, 701, 704-716, 719-738, 740-757, 759-766, 768-784, 786, 789-796, 800, 801, 803-812, 814-834, 836-865, 867-926, 928-1060, 1062-1094, 1096-1115, 1117-1126, 1129-1135, 1138-1144, 1146-1156, 1159-1165, 1168-1171, 1173-1175, 1177-1192, 1194-1196, 1198-1201, 1203, 1205-1227, 1229, 1232-1250, 1252, 1253, 1255-1258, 1260-1268, 1270-1272, 1274-1293, 1295 and 1296.

2. The nucleic acid molecule of claim 1, wherein said nucleic acid molecule is selected from the group consisting of RNA and DNA.

3. A vector comprising a nucleotide sequence selected from the group consisting of SEQ ID NOS: 4-6, 8, 12-14, 16-27, 30-50, 52, 53, 55-69, 72-86, 88-92, 97-100, 102-121, 123, 125, 127-144, 146-152, 154-164, 167, 168, 171-173, 175-182, 184-191, 193, 194, 196-205, 207, 209-214, 216-219, 221-225, 227-245, 250-252, 254, 256-260, 264-270, 273-281, 285-292, 294, 296-300, 303, 304, 306-313, 315-318, 320, 322-329, 331-335, 339-352, 354, 355, 357-388, 391, 393-426, 428, 429, 432-468, 470-486, 488-495, 497-507, 509-513, 515-525, 527, 528, 532, 533, 535, 538, 540-542, 545-556, 558-563, 565-576, 578-597, 599-618, 620, 624-628, 630, 631, 637-656, 659-662, 665, 667-675, 677-683, 685-697, 701, 704-716, 719-738, 740-757, 759-766, 768-

784, 786, 789-796, 800, 801, 803-812, 814-834, 836-865, 867-926, 928-1060, 1062-1094, 1096-1115, 1117-1126, 1129-1135, 1138-1144, 1146-1156, 1159-1165, 1168-1171, 1173-1175, 1177-1192, 1194-1196, 1198-1201, 1203, 1205-1227, 1229, 1232-1250, 1252, 1253, 1255-1258, 1260-1268, 1270-1272, 1274-1293, 1295 and 1296.

4. The vector of claim 3 wherein the vector is selected from the group consisting of a virus and a plasmid.

5. The vector of claim 4 wherein the nucleotide sequence is operably linked to a promoter sequence.

6. A host cell transformed with an expression vector comprising a nucleotide sequence selected from the group consisting of SEQ ID NOS: 4-6, 8, 12-14, 16-27, 30-50, 52, 53, 55-69, 72-86, 88-92, 97-100, 102-121, 123, 125, 127-144, 146-152, 154-164, 167, 168, 171-173, 175-182, 184-191, 193, 194, 196-205, 207, 209-214, 216-219, 221-225, 227-245, 250-252, 254, 256-260, 264-270, 273-281, 285-292, 294, 296-300, 303, 304, 306-313, 315-318, 320, 322-329, 331-335, 339-352, 354, 355, 357-388, 391, 393-426, 428, 429, 432-468, 470-486, 488-495, 497-507, 509-513, 515-525, 527, 528, 532, 533, 535, 538, 540-542, 545-556, 558-563, 565-576, 578-597, 599-618, 620, 624-628, 630, 631, 637-656, 659-662, 665, 667-675, 677-683, 685-697, 701, 704-716, 719-738, 740-757, 759-766, 768-784, 786, 789-796, 800, 801, 803-812, 814-834, 836-865, 867-926, 928-1060, 1062-1094, 1096-1115, 1117-1126, 1129-1135, 1138-1144, 1146-1156, 1159-1165, 1168-1171, 1173-1175, 1177-1192, 1194-1196, 1198-1201, 1203, 1205-1227, 1229, 1232-1250, 1252, 1253, 1255-1258, 1260-1268, 1270-1272, 1274-1293, 1295 and 1296.

7. A polypeptide encoded by a nucleotide sequence selected from the group consisting of SEQ ID NOS: 4-6, 8, 12-14, 16-27, 30-50, 52, 53, 55-69, 72-86, 88-92, 97-100, 102-121, 123, 125, 127-144, 146-152, 154-164, 167, 168, 171-173, 175-182, 184-191, 193, 194,

196-205, 207, 209-214, 216-219, 221-225, 227-245, 250-252, 254, 256-260, 264-270, 273-281, 285-292, 294, 296-300, 303, 304, 306-313, 315-318, 320, 322-329, 331-335, 339-352, 354, 355, 357-388, 391, 393-426, 428, 429, 432-468, 470-486, 488-495, 497-507, 509-513, 515-525, 527, 528, 532, 533, 535, 538, 540-542, 545-556, 558-563, 565-576, 578-597, 599-618, 620, 624-628, 630, 631, 637-656, 659-662, 665, 667-675, 677-683, 685-697, 701, 704-716, 719-738, 740-757, 759-766, 768-784, 786, 789-796, 800, 801, 803-812, 814-834, 836-865, 867-926, 928-1060, 1062-1094, 1096-1115, 1117-1126, 1129-1135, 1138-1144, 1146-1156, 1159-1165, 1168-1171, 1173-1175, 1177-1192, 1194-1196, 1198-1201, 1203, 1205-1227, 1229, 1232-1250, 1252, 1253, 1255-1258, 1260-1268, 1270-1272, 1274-1293, 1295 and 1296.

8. A polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NOS: 1300-1302, 1302, 1308-1310, 1312-1323, 1326-1346, 1348, 1349, 1351-1365, 1368-1382, 1384-1388, 1393-1396, 1398-1417, 1419, 1421, 1423-1440, 1442-1448, 1450-1460, 1463, 1464, 1467-1469, 1471-1478, 1480-1487, 1489, 1490, 1492-1501, 1503, 1505-1510, 1512-1515, 1517-1521, 1523-1541, 1546-1548, 1550, 1552-1556, 1560-1566, 1569-1577, 1581-1588, 1590, 1592-1596, 1599, 1600, 1602-1609, 1611-1614, 1616, 1618-1625, 1627-1631, 1635-1648, 1650, 1651, 1653-1684, 1687, 1689-1722, 1724, 1725, 1728-1764, 1766-1782, 1784-1791, 1793-1803, 1805-1809, 1811-1821, 1823, 1824, 1828, 1829, 1831, 1834, 1836-1838, 1841-1852, 1854-1859, 1861-1872, 1874-1893, 1895-1914, 1916, 1920-1924, 1926, 1927, 1933-1952, 1955-1958, 1961, 1963-1971, 1973-1979, 1981-1993, 1997, 2000-2012, 2015-2034, 2036-2053, 2055-2062, 2064-2080, 2082, 2085-2092, 2096, 2097, 2099-2108, 2110-2130, 2132-2161, 2163-2222, 2224-2356, 2358-2390, 2392-2411, 2413-2422, 2425-2431, 2434-2440, 2442-2452, 2455-2461, 2464-2467, 2469-2471, 2473-2488, 2490-2492, 2494-2497, 2499, 2501-2523, 2525, 2528-2546, 2548, 2549, 2551-2554, 2556-2564, 2568-2570, 2572-2589, 2591 and 2592.

9. An isolated antibody that is capable of binding to a polypeptide encoded by nucleotide sequence selected from the group consisting of SEQ ID NOS: 4-6, 8, 12-14, 16-27, 30-50, 52, 53, 55-69, 72-86, 88-92, 97-100, 102-121, 123, 125, 127-144, 146-152, 154-164, 167, 168, 171-173, 175-182, 184-191, 193, 194, 196-205, 207, 209-214, 216-219, 221-225, 227-245, 250-252, 254, 256-260, 264-270, 273-281, 285-292, 294, 296-300, 303, 304, 306-313, 315-318, 320, 322-329, 331-335, 339-352, 354, 355, 357-388, 391, 393-426, 428, 429, 432-468, 470-486, 488-495, 497-507, 509-513, 515-525, 527, 528, 532, 533, 535, 538, 540-542, 545-556, 558-563, 565-576, 578-597, 599-618, 620, 624-628, 630, 631, 637-656, 659-662, 665, 667-675, 677-683, 685-697, 701, 704-716, 719-738, 740-757, 759-766, 768-784, 786, 789-796, 800, 801, 803-812, 814-834, 836-865, 867-926, 928-1060, 1062-1094, 1096-1115, 1117-1126, 1129-1135, 1138-1144, 1146-1156, 1159-1165, 1168-1171, 1173-1175, 1177-1192, 1194-1196, 1198-1201, 1203, 1205-1227, 1229, 1232-1250, 1252, 1253, 1255-1258, 1260-1268, 1270-1272, 1274-1293, 1295 and 1296.

10. A method for identifying compounds useful for treating anosmia comprising (a) identifying a candidate compound by its ability to bind to a polypeptide encoded by nucleotide sequence selected from the group consisting of SEQ ID NOS: 4-6, 8, 12-14, 16-27, 30-50, 52, 53, 55-69, 72-86, 88-92, 97-100, 102-121, 123, 125, 127-144, 146-152, 154-164, 167, 168, 171-173, 175-182, 184-191, 193, 194, 196-205, 207, 209-214, 216-219, 221-225, 227-245, 250-252, 254, 256-260, 264-270, 273-281, 285-292, 294, 296-300, 303, 304, 306-313, 315-318, 320, 322-329, 331-335, 339-352, 354, 355, 357-388, 391, 393-426, 428, 429, 432-468, 470-486, 488-495, 497-507, 509-513, 515-525, 527, 528, 532, 533, 535, 538, 540-542, 545-556, 558-563, 565-576, 578-597, 599-618, 620, 624-628, 630, 631, 637-656, 659-662, 665, 667-675, 677-683, 685-697, 701, 704-716, 719-738, 740-757, 759-766, 768-784, 786, 789-796, 800, 801, 803-812, 814-834,

836-865, 867-926, 928-1060, 1062-1094, 1096-1115, 1117-1126, 1129-1135, 1138-1144, 1146-1156, 1159-1165, 1168-1171, 1173-1175, 1177-1192, 1194-1196, 1198-1201, 1203, 1205-1227, 1229, 1232-1250, 1252, 1253, 1255-1258, 1260-1268, 1270-1272, 1274-1293, 1295 and 1296; (b) introducing the candidate compound to an anosmic mouse; and (c) determining whether the mouse detects isovaleric acid at a concentration less than  $10^{-5}$  M after the candidate compound has been introduced, wherein the ability of the mouse to detect isovaleric acid at a concentration less than  $10^{-5}$  M indicates that the compound is useful for the treatment of anosmia.

11. A method for identifying a compound for pest control comprising (a) identifying a candidate compound by its ability to bind to a polypeptide encoded by nucleotide sequence selected from the group consisting of SEQ ID NOS: 4-6, 8, 12-14, 16-27, 30-50, 52, 53, 55-69, 72-86, 88-92, 97-100, 102-121, 123, 125, 127-144, 146-152, 154-164, 167, 168, 171-173, 175-182, 184-191, 193, 194, 196-205, 207, 209-214, 216-219, 221-225, 227-245, 250-252, 254, 256-260, 264-270, 273-281, 285-292, 294, 296-300, 303, 304, 306-313, 315-318, 320, 322-329, 331-335, 339-352, 354, 355, 357-388, 391, 393-426, 428, 429, 432-468, 470-486, 488-495, 497-507, 509-513, 515-525, 527, 528, 532, 533, 535, 538, 540-542, 545-556, 558-563, 565-576, 578-597, 599-618, 620, 624-628, 630, 631, 637-656, 659-662, 665, 667-675, 677-683, 685-697, 701, 704-716, 719-738, 740-757, 759-766, 768-784, 786, 789-796, 800, 801, 803-812, 814-834, 836-865, 867-926, 928-1060, 1062-1094, 1096-1115, 1117-1126, 1129-1135, 1138-1144, 1146-1156, 1159-1165, 1168-1171, 1173-1175, 1177-1192, 1194-1196, 1198-1201, 1203, 1205-1227, 1229, 1232-1250, 1252, 1253, 1255-1258, 1260-1268, 1270-1272, 1274-1293, 1295 and 1296; (b) introducing the candidate compound in the vicinity of an animal; and (c) examining the behavior of the animal to determine whether the animal moves towards the candidate compound wherein the movement towards the candidate compound indicates that the candidate compound attract

pests.

12. A method for identifying a compound for pest control comprising (a) identifying a candidate compound by its ability to bind to a polypeptide encoded by nucleotide sequence selected from the group consisting of SEQ ID NOS: 4-6, 8, 12-14, 16-27, 30-50, 52, 53, 55-69, 72-86, 88-92, 97-100, 102-121, 123, 125, 127-144, 146-152, 154-164, 167, 168, 171-173, 175-182, 184-191, 193, 194, 196-205, 207, 209-214, 216-219, 221-225, 227-245, 250-252, 254, 256-260, 264-270, 273-281, 285-292, 294, 296-300, 303, 304, 306-313, 315-318, 320, 322-329, 331-335, 339-352, 354, 355, 357-388, 391, 393-426, 428, 429, 432-468, 470-486, 488-495, 497-507, 509-513, 515-525, 527, 528, 532, 533, 535, 538, 540-542, 545-556, 558-563, 565-576, 578-597, 599-618, 620, 624-628, 630, 631, 637-656, 659-662, 665, 667-675, 677-683, 685-697, 701, 704-716, 719-738, 740-757, 759-766, 768-784, 786, 789-796, 800, 801, 803-812, 814-834, 836-865, 867-926, 928-1060, 1062-1094, 1096-1115, 1117-1126, 1129-1135, 1138-1144, 1146-1156, 1159-1165, 1168-1171, 1173-1175, 1177-1192, 1194-1196, 1198-1201, 1203, 1205-1227, 1229, 1232-1250, 1252, 1253, 1255-1258, 1260-1268, 1270-1272, 1274-1293, 1295 and 1296; (b) introducing the candidate compound in the vicinity of an animal; and (c) examining the behavior of the animal to determine whether the animal moves away from the candidate compound wherein the movement away from the candidate compound indicates that the candidate compound repels pests.

13. A method for identifying a compound which inhibits a mouse OR comprising (a) contacting a candidate compound with a mouse OR protein under condition that allow the mouse OR protein to interact with the target protein and (b) monitoring for contact between the mouse OR protein and the candidate compound using an assay wherein said mouse protein is encoded by a nucleotide sequence selected from the group consisting of SEQ NOS: 4-6, 8, 12-14, 16-27,

30-50, 52, 53, 55-69, 72-86, 88-92, 97-100, 102-121, 123, 125, 127-144, 146-152, 154-164, 167, 168, 171-173, 175-182, 184-191, 193, 194, 196-205, 207, 209-214, 216-219, 221-225, 227-245, 250-252, 254, 256-260, 264-270, 273-281, 285-292, 294, 296-300, 303, 304, 306-313, 315-318, 320, 322-329, 331-335, 339-352, 354, 355, 357-388, 391, 393-426, 428, 429, 432-468, 470-486, 488-495, 497-507, 509-513, 515-525, 527, 528, 532, 533, 535, 538, 540-542, 545-556, 558-563, 565-576, 578-597, 599-618, 620, 624-628, 630, 631, 637-656, 659-662, 665, 667-675, 677-683, 685-697, 701, 704-716, 719-738, 740-757, 759-766, 768-784, 786, 789-796, 800, 801, 803-812, 814-834, 836-865, 867-926, 928-1060, 1062-1094, 1096-1115, 1117-1126, 1129-1135, 1138-1144, 1146-1156, 1159-1165, 1168-1171, 1173-1175, 1177-1192, 1194-1196, 1198-1201, 1203, 1205-1227, 1229, 1232-1250, 1252, 1253, 1255-1258, 1260-1268, 1270-1272, 1274-1293, 1295 and 1296.

14. A method for identifying a compound that stimulates a mouse OR comprising (a) contacting a candidate compound with a mouse OR protein under condition that allow the mouse OR protein to interact with the target protein and (b) monitoring for contact between the mouse OR protein and the candidate compound using an assay wherein said mouse protein is encoded by a nucleotide sequence selected from the group consisting of SEQ NOS: 4-6, 8, 12-14, 16-27, 30-50, 52, 53, 55-69, 72-86, 88-92, 97-100, 102-121, 123, 125, 127-144, 146-152, 154-164, 167, 168, 171-173, 175-182, 184-191, 193, 194, 196-205, 207, 209-214, 216-219, 221-225, 227-245, 250-252, 254, 256-260, 264-270, 273-281, 285-292, 294, 296-300, 303, 304, 306-313, 315-318, 320, 322-329, 331-335, 339-352, 354, 355, 357-388, 391, 393-426, 428, 429, 432-468, 470-486, 488-495, 497-507, 509-513, 515-525, 527, 528, 532, 533, 535, 538, 540-542, 545-556, 558-563, 565-576, 578-597, 599-618, 620, 624-628, 630, 631, 637-656, 659-662, 665, 667-675, 677-683, 685-697, 701, 704-716, 719-738, 740-757, 759-766, 768-784, 786, 789-796, 800, 801, 803-812,

814-834, 836-865, 867-926, 928-1060, 1062-1094, 1096-1115, 1117-1126, 1129-1135, 1138-1144, 1146-1156, 1159-1165, 1168-1171, 1173-1175, 1177-1192, 1194-1196, 1198-1201, 1203, 1205-1227, 1229, 1232-1250, 1252, 1253, 1255-1258, 1260-1268, 1270-1272, 1274-1293, 1295 and 1296.

15. A data mining method for identifying gene family members comprising

(a) providing reference organism gene family sequences, candidate organism gene family sequences and query sequences;

(b) performing a low stringency TBLASTN search using the query sequence to obtain a pool of potential candidate sequences from a candidate organism database;

(c) building a profile Hidden Markov Model (HMM) using sequences from the reference and candidate organism to generate a profile HMM;

(d) comparing the pool of potential candidate sequences with the profile HMM to generate a pool of candidate sequences; and

(e) determining whether the candidate sequences obtained in step (d) are gene family members.

16. A method for identifying OR sequences comprising

(a) providing reference organism gene family sequences, candidate organism gene family sequences and query sequences;

(b) performing a low stringency TBLASTN search using the query sequences to obtain a pool of potential candidate OR sequences;

(c) building a profile Hidden Markov Model (HMM) using the OR sequences from a reference organism and a candidate organism to generate a profile HMM;

(d) comparing the pool of potential candidate OR sequences with the profile



HMM to generate a pool of candidate OR sequences; and

(e) determining whether the candidate sequences obtained in step (d) are OR gene family members.